The Mind Shift in Teacher Evaluation

Where We Stand—and Where We Need to Go

BY ANGELA MINNICI

The teaching profession today is full of contradictions: Teach the whole child, but focus on specific needs. Integrate 21st-century technology, yet get back to the basics. While often cited as one of the most rewarding professions, teaching is demanding, technically challenging, and more closely scrutinized by the public than ever. Although we better understand how children learn and how to support and develop educators, teachers today report more dissatisfaction with their jobs[1] and are less likely to stay in the profession beyond five years.[2]

The past decade has confirmed what we intuitively know—

Angela Minnici is a principal researcher in the Education Program at the American Institutes for Research, where she focuses on teacher evaluation and development. She also is the director of the Center on Great Teachers and Leaders, a federally funded program dedicated to advancing state efforts to grow, respect, and retain great teachers and leaders.

teachers are the most important in-school factor that influences student achievement.[3] Yet, at the same time, recent years have illuminated the field's struggle to unpack the nuances of teaching practices that have the greatest potential for improving student achievement. So how do we ensure *all* students are college- and career-ready in an era that is constantly evolving and in flux? How can we foster student achievement in all our schools? And how do we know what effective instruction looks like for each learner? These questions have prompted recent changes in federal policy and state legislation.

While much of the policy conversation about teachers over the last decade has focused on accountability, teaching quality is fundamentally an equity issue. Currently, federal, state, and local policymakers have advocated teacher evaluation systems as the solution to improving teaching quality and ultimately to addressing equity issues. But can teacher evaluation systems, as currently designed and implemented, improve teaching practices?[4] Can they help all teachers grow throughout their careers? What lessons learned must we incorporate to make these systems successful?

This article takes stock of teacher evaluation by exploring the successes and challenges of implementing evaluation systems. It also offers recommendations that my colleagues and I at the American Institutes for Research (AIR) have gleaned from our work with states and districts in designing and implementing systems that support and develop teachers.

Across the educator career continuum, AIR is conducting and applying high-quality, relevant research to support states, districts, and educators in promoting and sustaining teaching quality. For example, in one state, our team worked with the Department of Elementary and Secondary Education on all aspects of designing and implementing the state's model evaluation system, including training, developing tools and materials, and incorporating measures of student learning into evaluation. In another example, we collaborated with one district to create a new principal evaluation system that relies on multiple measures of performance and helps educators grow professionally through clear feedback and reflective dialogue. The new principal evaluation system included a standards-aligned and research-based evaluation framework, tools for principal supervisors to conduct observations of principals providing instructional feedback, and a school working-conditions survey on staff perceptions of school climate. In another district, our team enhanced educator support by integrating a redesigned career ladder with teacher leadership opportunities.

AIR is the lead partner in several federally funded content and regional technical assistance centers, which are a part of the Comprehensive Centers Program, that support states in their efforts to implement education reforms.* For example, the Center on Great Teachers and Leaders is dedicated to helping states grow, respect, and support great teachers and leaders for all students. Through our work in the center, we provide content-specific knowledge, expertise, and analyses to states and serve as a national resource on effective policies and practices to strengthen the quality of teaching and leading—especially in high-poverty, low-performing, and hard-to-staff schools.

In addition to helping states and districts implement programs, policies, and practices designed to improve teaching and leading, our team also designs and conducts implementation studies of educator evaluation, teacher and leader professional development, and mentoring and induction programs. At AIR, we believe in ensuring that rigorous research and evidence are used to address educational problems, and that policy decisions, in turn, are based on what is learned from research and evidence.

Early Successes

State teacher evaluation policies underwent sweeping changes in 2008 with the advent of federal competitive funding opportunities (e.g., Race to the Top and Investing in Innovation) in an economically challenging time. State legislative reforms continued as state leaders were offered the opportunity to obtain waivers from the No Child Left Behind Act in exchange for agreeing to reform how teachers and leaders are evaluated.[5] Currently, 49 states and the District of Columbia have changed their teacher evaluation legislation or guidance to reflect a fairly consistent vision of high-quality educator evaluation systems.[6] During the

past five years, states and districts have worked, sometimes at a feverish pace, to implement these changes, and tangible lessons have been learned.[7] We have worked with states and districts as they have created solutions to the dozens of technical and practical aspects of implementing evaluation systems. Yet, the most important lessons learned focus less on the technical aspects of the work and more on the fundamental mind shifts that have occurred and that ultimately have laid the foundation for this effort's success. These mind shifts, which have resulted in significant success in the field, are discussed below.

A common understanding about effective practice is the basis for high-quality evaluation systems that can drive professional growth.

First, among several critical components for effective implementation and sustainability of teacher evaluation systems, is the need to define and agree on what good teaching is. Teachers and administrators need a common language and vision about what constitutes effective practice. Clearly articulating these practices allows administrators to assess teachers and provide them with feedback on their strengths and areas for growth. It also encourages teachers and administrators to engage in professional conversations that make the critical link between teaching and the supports that teachers need to improve and hone their skills. This common understanding is the basis for high-quality evaluation systems that can drive professional growth.

Implicitly related to defining good teaching is defining the evidence and measures that are used to assess practice. Although more work is needed—particularly in ensuring a more balanced and evidence-based approach to combining and weighting measures[8]—states and districts have made considerable progress in identifying and employing measures that are more consistent and accurate. The conversation in many places has begun

to shift from ensuring only the validity and reliability of measures to a more nuanced conversation about the need for a variety of measures for a variety of reasons. For example, educators are asking questions about what measures and evidence will truly help them improve instruction, by providing them with information about teaching practices strongly linked to positive student outcomes.

It's also important to keep in mind that teacher engagement throughout the design and development process is not merely beneficial but critical to success. Teachers, as the experts in their craft, have much to contribute to the design and implementation of teacher evaluation systems. Their engagement throughout the process promotes ownership and efficacy of the system. These

Teachers, as experts in their craft, have much to contribute to the design and implementation of teacher evaluation systems.

systems are more likely to produce the results we desire—improved teaching quality and increased student learning—when teachers believe the systems and approaches will help them be more effective with their students.

Additionally, changes in evaluation processes and outcomes require not just new tools and procedures but significant changes in norms and beliefs. Teacher evaluation in most districts prior to 2008 was perfunctory and did little to help teachers improve.[9] In the early stages of this work, when state- and district-level committees were focused on designing teacher evaluation systems with all the necessary components, not enough thought and time was given (sometimes because of accelerated timelines) to the necessary culture shifts and new knowledge and skills at the district and school levels required to do this work.

These systems require teachers and administrators to think and act very differently in the evaluation process. It requires

them to make the shift from *sorting* to *supporting*. For example, new systems often ask administrators (and in some districts, teacher leaders) to hold evidence-driven conversations with teachers about instruction and student learning and to connect that evidence to district and school supports that will drive individual and school-wide improvement. Although states and districts are beginning to focus their attention and resources on these kinds of implementation challenges, much work remains in changing norms and beliefs.

Common Missteps to Avoid

To guide states and districts in designing and implementing effective evaluation systems, I've compiled a top 10 list of missteps to avoid, in order of importance. (For more specific examples of such missteps and how one school district avoided them, see the article on page 18.)

1. **Thinking teacher evaluation alone is the silver bullet.** Teacher evaluation systems alone are insufficient to improve instructional quality and increase student achievement. To be successful, reform efforts need to be coherent and aligned across the educator career continuum, beginning with recruitment and preparation, and extending to support, evaluation, and compensation.

2. **Excluding educators from the work.** Educators need to be extensively involved. This does not mean that only a handful of teachers on a state or district committee will suffice. Involvement must be broad and deep.

3. **Dismissing the importance of building trust.** Teachers need to believe that these systems will help support them and achieve success with their students. In many schools across the country, mistrust among educators exists for many reasons. The lack of trust at the district and school levels will likely affect the success of these systems, so take steps to begin rebuilding trust. The first step is ensuring that teachers and their unions are substantively involved in the design and implementation process.

4. **Failing to communicate frequently.** Regular communication is critical to the successful implementation of teacher evaluation systems. States and districts must develop communication plans that outline multiple ways to reach educators, parents, the community, and other important stakeholders about the effort. Communication also must be designed specifically to inform educators throughout the implementation process.

5. **Relying on principals to do all the work.** The role of the evaluator in these systems requires new skills and more time, and many principals have not received adequate training to carry out this new role. These systems will require new approaches to implementation, such as redistributing administrators' current work or rethinking staffing roles, including the role of teacher leaders in the evaluation process.

6. **Inadvertently decoupling teacher evaluation from professional learning.** This misstep occurs both in the

design and implementation phases. First, teacher evaluation systems have little chance of improving teaching if they do not include varied measures that can yield rich and relevant information about teacher practice.[10] Second, when teacher evaluation systems are implemented, administrators need to intentionally make the connections between evaluation information and professional learning opportunities.[11] Teachers need specific recommendations about and access to professional learning opportunities linked to evaluation outcomes.

7. **Underestimating time and resources.** Getting teacher evaluation right is a continuous process, not a one-time activity or event. It will require a significant investment of time and resources, particularly to identify evidence and measures of student learning to incorporate into the process that will help teachers improve their practice and to build the knowledge and skills of teachers and administrators. Make sure to lay the foundation for this work as a process of continuous improvement.

8. **Communicating the wrong message.** The wrong message about this work can derail well-intentioned efforts. Teachers and administrators are unlikely to want to put in the time and energy necessary to make these systems work if they think this effort is unimportant, unproven, and compliance driven.

9. **Not connecting the dots.** Do district policies feel like random acts of improvement? Many programs can support and guide high-quality instruction, such as mentoring and induction programs, instructional coaching, and professional learning communities. Take the time to figure out how all of these efforts can work together.

10. **Going it alone.** Considerable progress has been made in teacher evaluation design and implementation. States and districts do not have to reinvent the wheel as they begin this work. They should leverage their resources by partnering with other states and districts or using free resources such as those found on the Center on Great Teachers and Leaders website (www.gtlcenter.org), the Everyone at the Table website (www.everyoneatthetable.org), and NYSUT's Teacher Evaluation and Development website (www.nysut.org/resources/special-resources-sites/ted).

Despite making significant progress, states and districts still face challenges that could threaten the ultimate success of these systems. At AIR, we anticipate two major ones:

Inability to stay the course. Changes take time to implement. In education, we rarely allocate sufficient time for the implementation and careful study of any major policy change. We often rush to proclaim an effort unsuccessful and then quickly move on to usher in another one. Lessons learned from the early adopters in the field (e.g., Denver Public Schools, New Haven Public Schools, and Tennessee) suggest that we need at least five to seven years of implementation before we begin to assess the effectiveness of such efforts. And, given the sweeping changes occurring in many states and districts regarding the implementation of the Common Core State Standards, we might need even more time to decide whether to continue a particular teacher evaluation effort.

New ideas and processes require system change. Creating systems that can improve instructional quality will require changes in the way we organize and deliver schooling. For example, teachers will need more time to collaborate with peers and to study and reflect on their practice if we want them to develop and improve at all stages of their careers. Principals will need more time in classrooms to have conversations with teachers about effective practice. Principals and teachers will need the right set of supports (professional development and beyond) to deepen their knowledge about effective practice. We are designing new systems and ways of doing business regarding teacher evaluation, yet we continue to try to fit those new approaches into the traditional model of schooling. If we are to observe real improvements in teaching and learning, then we will likely need to restructure the school day and reallocate existing resources so that teachers have more time for studying and improving their practice with colleagues. School days must be designed with teacher collaboration in mind, instead of trying to cram collaboration into the school day.

As we look at how far we have come and how far we need to go in teacher evaluation implementation, states and districts can take a few key steps to increase the likelihood that teacher evaluation systems will help improve instructional quality on a broad and deep scale.

Real improvements in teaching and learning come when teachers have more time for studying and improving their practice with colleagues.

Do not rush these reforms. Balance the political need to move forward and show progress with the pragmatic consideration of making sure these systems are doing what they are intended to do—support and develop educators so that their students are academically, socially, and emotionally successful. Too much time, effort, and resources have already been spent on teacher evaluation to let these systems fail because we do not have the patience to see these efforts through. Teachers and administrators need time to learn these new systems and to gain the trust and confidence that they will actually support rather than merely sort teachers. Finally, expecting to observe real outcomes in student learning that can be attributed to teacher evaluation systems will take time and close study of the ways in which these systems are improving instructional quality.

Teacher evaluation by itself is inefficient for improving the quality of all teachers.

Keep educators substantively engaged. Implementation is messy, and no state or district gets this correct right out of the gate. If educators are not substantively involved in the design and implementation of these systems, or they do not believe these systems will actually support them, then they are less likely to want to stay the course and invest the time and effort needed to fix the problems that will most definitely occur during implementation. Disinvestment and mistrust of educators in this work will crumble the system's foundation quickly.

Put together all the pieces of the puzzle. As the article on page 4 makes clear, teacher evaluation is only one component of a systems approach to supporting, developing, and improving teaching quality in a state or district. Work with a broad group of stakeholders to map out how other important initiatives (e.g., induction and mentoring programs, peer assistance and review, lesson study, professional learning communities, and response to intervention) also support the goal of high-quality teaching for all students. Help educators see clearly how these programs and initiatives all work together to support high-quality instruction.

Although much of this article has focused on teacher evaluation and its promise to improve teaching quality, it is important to point out that teacher evaluation by itself is an inefficient approach to significantly improving the quality of all teachers. Who we recruit into the profession and how we prepare them are just as essential as how we develop, support, and retain them once they enter the classroom. States and districts need to develop a coherent, comprehensive, and coordinated approach to improving teaching quality. To ensure educational equity, such an approach must include teacher evaluation systems designed to help all teachers develop and improve throughout their careers. ☐

Endnotes

1. *The MetLife Survey of the American Teacher: Challenges for School Leadership* (New York: MetLife, 2012), www.metlife.com/assets/cao/foundation/MetLife-Teacher-Survey-2012.pdf.

2. Richard M. Ingersoll and Lisa Merrill, *Seven Trends: The Transformation of the Teaching Force*, updated October 2013 (Philadelphia: Consortium for Policy Research in Education, 2013), www.cpre.org/sites/default/files/workingpapers/1506_seventrendsupdated october2013.pdf.

3. Barbara Nye, Spyros Konstantopoulos, and Larry V. Hedges, "How Large Are Teacher Effects?" *Educational Evaluation and Policy Analysis* 26 (2004): 237–257.

4. Sixteen states have passed *legislation* explicitly stating that the purpose of teacher evaluations is to improve teacher practice and/or promote professional growth. Forty-two states have *guidance documents* explicitly stating that the purpose of teacher evaluations is to improve teacher practice and/or promote professional growth. See Center on Great Teachers and Leaders, *Databases on State Teacher and Principal Evaluation Policies* (2013), http://resource.tqsource.org/stateevaldb.

5. To date, 42 states and the District of Columbia have been approved for ESEA flexibility. See U.S. Department of Education, "ESEA Flexibility," updated on December 13, 2013, www2.ed.gov/policy/elsec/guid/esea-flexibility/index.html.

6. Center on Great Teachers and Leaders, *Database on State Teacher and Principal Evaluation Policies*.

7. See, for example, Bill & Melinda Gates Foundation, *Ensuring Fair and Reliable Measures of Effective Teaching: Culminating Findings from the MET Project's Three-Year Study* (Seattle: Bill & Melinda Gates Foundation, 2013); Corinne Herlihy, Ezra Karger, Cynthia Pollard, Heather C. Hill, Matthew A. Kraft, Megan Williams, and Sarah Howard, "State and Local Efforts to Investigate the Validity and Reliability of Scores from Teacher Evaluation Systems," *Teachers College Record* 116, no. 1 (2014); and Heather C. Hill and Pam Grossman, "Learning from Teacher Observations: Challenges and Opportunities Posed by New Teacher Evaluation Systems," *Harvard Educational Review* 82 (2013): 371–384.

8. Recent research shows that including multiple measures of teaching practice and employing a balanced approach to weighting measures produces more-reliable teacher effectiveness ratings. See Bill & Melinda Gates Foundation, *Ensuring Fair and Reliable Measures*.

9. Daniel Weisberg, Susan Sexton, Jennifer Mulhern, and David Keeling, *The Widget Effect: Our National Failure to Acknowledge and Act on Differences in Teacher Effectiveness* (Brooklyn, NY: The New Teacher Project, 2009).

10. According to Goe, Holdheide, and Miller, "Although evidence of teacher effectiveness can be demonstrated, in part, through student growth measures, such measures are limited in distinguishing evidence of instructional quality." Laura Goe, Lynn Holdheide, and Tricia Miller, *A Practical Guide to Designing Comprehensive Teacher Evaluation Systems* (Washington, DC: National Comprehensive Center for Teacher Quality, 2011), 20. The authors recommend including measures such as observation instruments, performance rubrics, portfolios, evidence binders, teacher self-assessments, and parent and student surveys in teacher evaluation systems designed to help teachers improve their practice.

11. According to Coggshall et al., based on what we know about teacher learning and evaluation, teacher evaluation systems when implemented should "help teachers and school leaders develop a common understanding of the contours of effective practice," "provide evidence-based feedback to teachers to help them reflect on and improve their practice," and "measure and account for teachers' learning and collaboration." Jane G. Coggshall, Claudette Rasmussen, Amy Colton, Jessica Milton, and Catherine Jacques, *Generating Teacher Effectiveness: The Role of Job-Embedded Professional Learning in Teacher Evaluation* (Washington, DC: National Comprehensive Center for Teacher Quality, 2012), 6.